Geophysical Journal International

Geophys. J. Int. (2011) **186,** 264–278 doi: 10.1111/j.1365-246X.2011.05027.x

Adjoint centroid-moment tensor inversions

YoungHee Kim,[1] Qinya Liu[2] and Jeroen Tromp[3]

[1] *Seismological Laboratory, California Institute of Technology, Pasadena,* CA 91125, *USA. E-mail: ykim@gps.caltech.edu*
[2] *Department of Physics, University of Toronto, Toronto,* M5S 1A7, *Ontario, Canada*
[3] *Department of Geosciences and Program in Applied & Computational Mathematics, Princeton University, Princeton,* NJ 08544, *USA*

Accepted 2011 March 28. Received 2011 February 26; in original form 2011 January 6

SUMMARY

We determine centroid-moment tensor (CMT) solutions by minimizing waveform differences between observed and simulated seismograms based on an adjoint method. Synthetic seismograms and Fréchet derivatives are calculated based on a spectral-element method. The non-linear adjoint CMT inversion algorithm requires three simulations for each iteration: one 'forward' simulation to obtain synthetics for the current source parameters, one 'adjoint' simulation which involves injecting time-reversed differences between observed and simulated seismograms as simultaneous virtual sources at each of the receivers, and an extra forward simulation to compute the step length in the conjugate-gradient direction. Whereas the vertical component of the adjoint wavefield reflects the radiation pattern near the centroid location, the components of the adjoint strain tensor capture the elements of the moment tensor. We use the method to determine adjoint CMT solutions for two representative southern California earthquakes using recent 3-D crustal model CVM-6.2. The adjoint CMT solutions are in good agreement with classical Hessian-based CMT solutions involving 3-D Green's functions. In general, adjoint CMT inversions require fewer numerical simulations than traditional Hessian-based inversions. This faster convergence holds promise for multiple moment-tensor and kinematic rupture inversions in 3-D earth models.

Key words: Inverse theory; Earthquake dynamics; Earthquake source observations; Computational seismology; Theoretical seismology.

1 INTRODUCTION

Estimating accurate earthquake source parameters is one of the most important tasks in modern seismology. Since the early 1980s, the global Centroid-Moment Tensor (CMT) project (www.globalCMT.org) routinely determines earthquake source parameters for all $M_w \geq \sim 5.5$ global earthquakes (e.g. Dziewonski *et al.* 1981; Ekström 2007). At regional scales, body and surface waves are used to provide source parameters for intermediate earthquakes (e.g. Wallace *et al.* 1981; Ritsema & Lay 1993). For larger earthquakes, finite-fault inversions attempt to map spatial and temporal distributions of slip (e.g. Olson & Aspel 1982; Hartzell & Heaton 1983; Hartzell *et al.* 1996; Ji *et al.* 2002).

In recent years, there have been efforts to image the rupture process based on backprojection techniques (e.g. Ishii *et al.* 2005; Walker *et al.* 2005) and the concept of time reversal (e.g. Larmat *et al.* 2006, 2008, 2010). The advantage of backprojection or time-reversal methods is that they require no prior knowledge of the source itself, and enhance signal-to-noise ratios by utilizing a large number of stations, thereby stacking coherent information in seismic data. Furthermore, such methods facilitate physical interpretations of complicated earthquake rupture processes (e.g. Hjörleifsdóttir *et al.* 2009). Connections between time-reversal source imaging and conventional moment-tensor inversions are discussed by Tromp *et al.* (2005) and Kawakatsu & Montagner (2008).

Southern California is of particular interest to geoscientists due to its geological and tectonic complexity, and potential seismic hazards from active faults. The abundance of high-quality data from the Southern California Seismic Network (SCSN) has made it possible to study space–time distributions and source mechanisms of southern California earthquakes in great detail (e.g. Zhao & Helmberger 1994; Chen *et al.* 2005, 2010). Several source parameter inversion methods, utilizing different portions of three-component broad-band seismograms (e.g. P_{nl}, long-period surface waves, or both), have been applied to southern California earthquakes (e.g. Helmberger & Engen 1980; Dreger & Helmberger 1991; Thio & Kanamori 1995; Zhu & Helmberger 1996; Liu *et al.* 2004; Tan 2006; Hauksson *et al.* 2008).

Rapid recent advances in computer technology and numerical methods have enabled accurate simulations of seismic wave propagation in complex 3-D earth models both on regional (Komatitsch & Tromp 1999; Komatitsch *et al.* 2004; Olsen *et al.* 2006) and global scales (Komatitsch & Tromp 2002a,b). For southern California earthquakes, utilizing the CVM-H model developed by the Southern California

Earthquake Center (SCEC, Hauksson 2000; Zhu & Kanamori 2000; Süss & Shaw 2003; Lovely *et al.* 2006; Plesch *et al.* 2009; Lin *et al.* 2010), 3-D simulations now provide more accurate Green's functions and achieve significantly better waveform fits to observed seismograms (Komatitsch *et al.* 2004; Tape *et al.* 2009) than standard 1-D models (Hadley & Kanamori 1977). Taking advantage of the CVM-H model, Liu *et al.* (2004) determined source mechanisms for southern California earthquakes by computing synthetic seismograms and source-parameter Fréchet derivatives based on a spectral-element method (SEM).

Recently, Tape *et al.* (2009, 2010) determined a model of the southern California crust based on an iterative tomographic inversion utilizing an adjoint method for the calculation of structural Fréchet derivatives. The inversion used the SCEC CVM-H model as a starting model, utilized selected waveforms based on an automated window selection algorithm (Maggi *et al.* 2009), computed Fréchet derivatives based on the interaction between forward and adjoint wavefields and iteratively improved the model based on a non-linear conjugate–gradient algorithm. The updated model, CVM-H6.2, exhibits ± 30 per cent variations in shear wave speed relative to the CVM-H starting model, significantly reduces delay times, and generally improves waveform fits between data and synthetics.

Strong 3-D heterogeneity introduces significant systematic errors in source-parameter inversions based on simple 1-D models. Therefore, it is important to use 3-D Green's functions in point-source or finite-fault inversions for events in regions with complicated crustal structure, such as southern California. Unfortunately, numerical simulations of 3-D Green's functions are computationally expensive, rendering classical Hessian-based source inversions impractical.

In this study, we introduce the theoretical framework for adjoint centroid-moment tensor inversions based on a conjugate–gradient algorithm, particularly in the context of 3-D SEM simulations. We examine both vertical component adjoint displacements and adjoint strain tensors in the vicinity of the source for all the basic earthquake mechanisms, thereby illustrating relationships between mechanism and adjoint wavefield. We use the new southern California crustal model CVM-H6.2 to illustrate the inversion technique for two earthquakes, namely, the 2001 September 9, $M_w = 4.3$ Hollywood event and the 2008 July 29, $M_w = 5.4$ Chino Hills event, and compare our results with solutions obtained based on classical Hessian-based moment-tensor inversions (Liu *et al.* 2004). Finally, we also test robustness of the technique and investigate the importance of the initial solution and the choice of source parametrization.

2 ADJOINT CENTROID-MOMENT TENSOR INVERSION

Fréchet derivatives of misfit functions with respect to earthquake source parameters were derived previously for point-source (e.g. Tromp *et al.* 2005; Liu & Tromp 2006; Tape *et al.* 2007) and finite-fault inversions (e.g. Hjörleifsdóttir 2007) based either on a Lagrange multiplier method or source–receiver reciprocity. Here we introduce an alternative derivation in which the adjoint wavefield is naturally defined.

2.1 Fréchet derivatives

A CMT solution consists of a moment tensor $\mathbf{M}$, a centroid location $\mathbf{x}_s$ and a source-time function S. To improve the CMT source parameters of an earthquake, we seek to minimize squared differences between observed seismograms $\mathbf{d}$ at various receiver locations, $\mathbf{x}_r, r = 1, \ldots, N_r$, and corresponding synthetic seismograms $\mathbf{s}$ calculated for the current source model $\mathbf{m} = \{\mathbf{M}, \mathbf{x}_s, S\}$.

$$\chi = \frac{1}{2} \sum_{rp} \int w_{rp}(t) \, || \, \mathbf{s}(\mathbf{x}_r, t; \mathbf{m}) - \mathbf{d}(\mathbf{x}_r, t) \, ||^2 \, \mathrm{d}t, \tag{1}$$

where we note that data and synthetics may have been filtered, and where $w_{rp}(t)$ denotes a combined windowing function and weighting factor assigned to the pth window at the rth receiver. Taking the variation of χ gives

$$\delta\chi = \sum_{rp} \int w_{rp}(t) \, [\mathbf{s}(\mathbf{x}_r, t; \mathbf{m}) - \mathbf{d}(\mathbf{x}_r, t)] \cdot \delta\mathbf{s}(\mathbf{x}_r, t; \mathbf{m}) \, \mathrm{d}t. \tag{2}$$

For brevity, we omit some of the functional dependences, and recognizing that both $\mathbf{d}$ and $\mathbf{s}$ are real vector fields, transform expression (2) into the frequency domain based on Parseval's theorem:

$$\delta\chi = \frac{1}{2\pi} \sum_{rp} \int [w_{rp} \, (\mathbf{s} - \mathbf{d})]^*(\omega) \cdot \delta\mathbf{s}(\omega) \, \mathrm{d}\omega. \tag{3}$$

An asterisk $*$ denotes complex conjugation, and $[w_{rp}(\mathbf{s} - \mathbf{d})]^*(\omega)$ may be interpreted in the time domain as the time-reversed, filtered, time-windowed and weighted waveform difference between data and synthetics.

We relate variations in the misfit function χ to variations in the source parameters $\mathbf{m}$ via the seismic wave equation

$$\rho \, \ddot{\mathbf{s}} = \nabla \cdot (\mathbf{c} : \nabla\mathbf{s}) - \mathbf{M} \cdot \nabla\delta(\mathbf{x} - \mathbf{x}_s) \, S(t), \tag{4}$$

where a dot denotes partial differentiation with respect to time, ∇ the gradient with respect to $\mathbf{x}$, ρ and $\mathbf{c}$ mass density and the elastic tensor, respectively, and $\mathbf{M}$ the centroid-moment tensor representing the earthquake with centroid location $\mathbf{x}_s$. The source-time function is denoted by S, such that $\dot{S}$ is the moment-rate function. The initial displacement and velocity are assumed to vanish.

We linearize the inverse problem by taking the variation of eq. (4). Ignoring variations due to structural parameters, in the frequency domain, we obtain

$$-\omega^2 \rho \, \delta\mathbf{s} = \nabla \cdot (\mathbf{c} : \nabla\delta\mathbf{s}) - [\delta\mathbf{M} \cdot \nabla\delta(\mathbf{x} - \mathbf{x}_s) + \mathbf{M} \cdot \nabla\nabla_s\delta(\mathbf{x} - \mathbf{x}_s) \cdot \delta\mathbf{x}_s] \, S - \mathbf{M} \cdot \nabla\delta(\mathbf{x} - \mathbf{x}_s) \, \delta S, \tag{5}$$

where ∇_s denotes the gradient with respect to the centroid location $\mathbf{x}_s$. Assuming that the simulations start at time $t = -t_0$ and end at time $t = T - t_0$, where T denotes the total record length, we take the dot product of eq. (5) with the adjoint field $\mathbf{s}^\dagger(\omega)e^{i\omega(T-2t_0)}$—the meaning of which is yet to be revealed—integrating over both angular frequency and space:

$$
\begin{aligned}
-\frac{1}{2\pi}\iint \rho\,\omega^2\,\mathbf{s}^\dagger \mathrm{e}^{\mathrm{i}\omega(T-2t_0)}\cdot\delta\mathbf{s}\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x} &= \frac{1}{2\pi}\iint \mathbf{s}^\dagger \mathrm{e}^{\mathrm{i}\omega(T-2t_0)}\cdot[\nabla\cdot(\mathbf{c}:\nabla\delta\mathbf{s})]\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}\\
&-\frac{1}{2\pi}\iint \mathbf{s}^\dagger \mathrm{e}^{\mathrm{i}\omega(T-2t_0)}\cdot[\delta\mathbf{M}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s)+\mathbf{M}\cdot\nabla\nabla_s\delta(\mathbf{x}-\mathbf{x}_s)\cdot\delta\mathbf{x}_s]\,S\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}\\
&-\frac{1}{2\pi}\iint \mathbf{s}^\dagger \mathrm{e}^{\mathrm{i}\omega(T-2t_0)}\cdot[\mathbf{M}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s)]\,\delta S\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}. \qquad (6)
\end{aligned}
$$

Integration by parts reduces the previous expression to

$$
\begin{aligned}
&\frac{1}{2\pi}\iint[-\rho\,\omega^2\,\mathbf{s}^\dagger-\nabla\cdot(\mathbf{c}:\nabla\mathbf{s}^\dagger)]\mathrm{e}^{\mathrm{i}\omega(T-2t_0)}\cdot\delta\mathbf{s}\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}\\
&\quad=-\frac{1}{2\pi}\iint \mathbf{s}^\dagger \mathrm{e}^{\mathrm{i}\omega(T-2t_0)}\cdot[\delta\mathbf{M}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s)+\mathbf{M}\cdot\nabla\nabla_s\delta(\mathbf{x}-\mathbf{x}_s)\cdot\delta\mathbf{x}_s]\,S\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}\\
&\qquad-\frac{1}{2\pi}\iint \mathbf{s}^\dagger \mathrm{e}^{\mathrm{i}\omega(T-2t_0)}\cdot[\mathbf{M}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s)]\,\delta S\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}. \qquad (7)
\end{aligned}
$$

Upon comparing the left-hand side of the last expression with the right-hand side of eq. (3), we let the adjoint field $\mathbf{s}^\dagger$ be governed by the adjoint wave equation

$$
-\rho\,\omega^2\,\mathbf{s}^\dagger-\nabla\cdot(\mathbf{c}:\nabla\mathbf{s}^\dagger)=\sum_{rp}[w_{rp}\,(\mathbf{s}-\mathbf{d})]^*(\omega)\,\mathrm{e}^{-\mathrm{i}\omega(T-2t_0)}\delta(\mathbf{x}-\mathbf{x}_r). \qquad (8)
$$

Based on the fact that if the Fourier transform of $f(t)$ is $F(\omega)$, then the Fourier transform of $f(-t)$ is $F^*(\omega)$ and the Fourier transform of $f(T-t)$ is $F^*(\omega)\,\mathrm{e}^{-\mathrm{i}\omega T}$, we may rewrite the adjoint eq. (8) in the time domain as

$$
\rho\,\ddot{\mathbf{s}}^\dagger=\nabla\cdot(\mathbf{c}:\nabla\mathbf{s}^\dagger)+\sum_{rp}[w_{rp}\,(\mathbf{s}-\mathbf{d})](T-2t_0-t)\,\delta(\mathbf{x}-\mathbf{x}_r), \qquad (9)
$$

that is, the adjoint field $\mathbf{s}^\dagger$ is generated by using time-reversed, windowed, weighted and filtered waveform differences between data and synthetics as simultaneous virtual sources at the receiver locations. Note that, in general, calculation of forward and adjoint wavefields requires full 3-D numerical simulations. Equating the left-hand side of eq. (3) with the right-hand side of eq. (7) gives

$$
\begin{aligned}
\delta\chi\;=&-\frac{1}{2\pi}\iint \mathbf{s}^\dagger\cdot[\delta\mathbf{M}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s)+\mathbf{M}\cdot\nabla\nabla_s\delta(\mathbf{x}-\mathbf{x}_s)\cdot\delta\mathbf{x}_s]\,[S^*\,\mathrm{e}^{-\mathrm{i}\omega(T-2t_0)}]^*\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}\\
&-\frac{1}{2\pi}\iint \mathbf{s}^\dagger\cdot[\mathbf{M}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_s)]\,[\delta S^*\,\mathrm{e}^{-\mathrm{i}\omega(T-2t_0)}]^*\,\mathrm{d}\omega\,\mathrm{d}^3\mathbf{x}, \qquad (10)
\end{aligned}
$$

or in the time domain

$$
\begin{aligned}
\delta\chi\;=&\;\delta\mathbf{M}:\int\boldsymbol{\epsilon}^\dagger(\mathbf{x}_s,t)\,S(T-2t_0-t)\,\mathrm{d}t+\delta\mathbf{x}_s\cdot\nabla_s\int\mathbf{M}:\boldsymbol{\epsilon}^\dagger(\mathbf{x}_s,t)\,S(T-2t_0-t)\,\mathrm{d}t\\
&+\int\mathbf{M}:\boldsymbol{\epsilon}^\dagger(\mathbf{x}_s,t)\,\delta S(T-2t_0-t)\,\mathrm{d}t, \qquad (11)
\end{aligned}
$$

where $\boldsymbol{\epsilon}^\dagger=\frac{1}{2}[\nabla\mathbf{s}^\dagger+(\nabla\mathbf{s}^\dagger)^T]$ denotes the strain tensor associated with the adjoint wavefield $\mathbf{s}^\dagger$ and where we have invoked the symmetry of moment tensor $\mathbf{M}$.

Let us first focus on the Fréchet derivatives with respect to moment tensor $\mathbf{M}$ and centroid location $\mathbf{x}_s$. In components, we have

$$
\begin{aligned}
\frac{\partial\chi}{\partial M_{ij}}&=\int\epsilon^\dagger_{ij}(\mathbf{x}_s,t)\,S(T-2t_0-t)\,\mathrm{d}t,\\
\frac{\partial\chi}{\partial x_i^s}&=\int\partial_{x_i^s}[\mathbf{M}:\boldsymbol{\epsilon}^\dagger(\mathbf{x}_s,t)]\,S(T-2t_0-t)\,\mathrm{d}t. \qquad (12)
\end{aligned}
$$

We may rewrite expressions (12) in terms of the moment-rate function $\dot{S}$ as

$$
\begin{aligned}
\frac{\partial\chi}{\partial M_{ij}}&=\int I\left[\epsilon^\dagger_{ij}(\mathbf{x}_s,t)\right]\,\dot{S}(T-2t_0-t)\,\mathrm{d}t,\\
\frac{\partial\chi}{\partial x_i^s}&=\int I\left[\partial_{x_i^s}(\mathbf{M}:\boldsymbol{\epsilon}^\dagger(\mathbf{x}_s,t))\right]\,\dot{S}(T-2t_0-t)\,\mathrm{d}t, \qquad (13)
\end{aligned}
$$

where $I[\cdot]$ denotes the time-integration operation. Since the moment-rate function $\dot{S}$ is dominated by energy released at the start of the earthquake, the integrated quantities $I[\epsilon^\dagger_{ij}]$ and $I[\partial_{x_i^s}(\mathbf{M}:\boldsymbol{\epsilon}^\dagger)]$ only contribute to the gradient towards the end of the adjoint simulation. In practice, adjoint SEM simulations are launched by injecting the time-reversed adjoint source $\sum_{rp}[w_{rp}\,(\mathbf{s}-\mathbf{d})](T-2t_0-t)\,\delta(\mathbf{x}-\mathbf{x}_r)$ simultaneously at all receivers. At each time step, the adjoint wavefield and strain tensor are extracted at the original source location $\mathbf{x}_s$ and multiplied by the source-time function S on the fly based on eq. (12). At the end of the adjoint simulation, $\partial\chi/\partial M_{ij}$ and $\partial\chi/\partial x_i^s$ may thus be obtained and stored. Alternatively, the time-integrated functions $I[\epsilon^\dagger_{ij}]$ and $I[\partial_{x_i^s}(\mathbf{M}:\boldsymbol{\epsilon}^\dagger)]$ may be saved as time-series at $\mathbf{x}_s$ and convolved

with the moment-rate function $\dot{S}$ afterwards, as in eq. (13). In this approach, the integrated adjoint strain due to a delta moment-rate function may be calculated and stored during numerical simulations, making it flexible to convolve with any arbitrary moment-rate function $\dot{S}$ during post-processing. Note from eqs (12) and (13) that when the adjoint simulation launches at time $t = -t_0$, the values of the gradient should be extracted at time $t = T - 2t_0$.

Next, let us consider perturbations in the source-time function, δS. We have

$$\delta\chi = \int \mathbf{M} : \boldsymbol{\epsilon}^\dagger(\mathbf{x}_s, t)\, \delta S(T - 2t_0 - t)\, \mathrm{d}t = \int I[\mathbf{M} : \boldsymbol{\epsilon}^\dagger(\mathbf{x}_s, t)]\, \delta \dot{S}(T - 2t_0 - t)\, \mathrm{d}t. \tag{14}$$

Suppose we parametrize the moment-rate function in terms of a centroid time t_s and a half duration h_s based on a Gaussian (Komatitsch & Tromp (2002a) demonstrate that a Gaussian moment-rate function closely resembles a triangle)

$$\dot{S} = \frac{1}{\sqrt{\pi}\, h_s}\, \mathrm{e}^{-\left(\frac{t - t_s}{h_s}\right)^2}. \tag{15}$$

Then

$$\frac{\partial S}{\partial t_s} = -\dot{S}, \tag{16}$$

$$\frac{\partial S}{\partial h_s} = -\left(\frac{t - t_s}{h_s}\right) \dot{S}, \tag{17}$$

and the corresponding Fréchet derivatives are

$$\frac{\partial \chi}{\partial t_s} = \int \mathbf{M} : \boldsymbol{\epsilon}^\dagger(\mathbf{x}_s, t)\, \partial_{t_s} S\,(T - 2t_0 - t)\, \mathrm{d}t, \tag{18}$$

$$\frac{\partial \chi}{\partial h_s} = \int \mathbf{M} : \boldsymbol{\epsilon}^\dagger(\mathbf{x}_s, t)\, \partial_{h_s} S\,(T - 2t_0 - t)\, \mathrm{d}t. \tag{19}$$

When multiple centroid-moment tensors are used to capture a compound or kinematic source, each subsource may be represented by a Gaussian moment-rate function, as in eq. (15), for $s = 1, \ldots, N_s$, with corresponding centroid-time and half-duration Fréchet derivatives given by eqs (18) and (19).

2.1.1 Hessian-based inversion

We have thus far discussed the computation of first-order derivatives of the misfit function based on adjoint methods. This already involves one forward and one adjoint simulation, which may be numerically demanding for 3-D earth models. If second-order derivatives are computed to facilitate Hessian-based inversions, we need to determine the Hessian matrix

$$H_{ij} = \frac{\partial^2 \chi}{\partial m_i\, \partial m_j} \sim \sum_{rp} \int w_{rp}(t)\, \partial_{m_i} \mathbf{s}(\mathbf{x}_r, t; \mathbf{m}) \cdot \partial_{m_j} \mathbf{s}(\mathbf{x}_r, t; \mathbf{m})\, \mathrm{d}t. \tag{20}$$

The implication is that we need to have access to derivatives at individual receiver locations, $\partial_{m_i} \mathbf{s}(\mathbf{x}_r, t; \mathbf{m})$. These derivatives can only be obtained through brute-force numerical differentiation, as in Liu *et al.* (2004), and the number of simulations involved is the product of the number of CMT source parameters (typically 6–10) with the number of subevents. This large number of simulations may be computationally prohibitive, especially when many subevents are involved. Therefore, even though access to the Hessian matrix generally results in faster convergence (Tape *et al.* 2007), it may be necessary to minimize the misfit function based solely on first-order derivatives using conjugate-gradient techniques.

2.2 Conjugate-gradient method

2.2.1 Scaling and pre-conditioning

It is important to non-dimensionalize the source model parameters based on a proper covariance matrix, as discussed by Tape *et al.* (2007). We assume a simple model covariance matrix of the form

$$\mathbf{C_m} = \begin{bmatrix} \sigma_1^2, & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_{N_m}^2 \end{bmatrix}, \tag{21}$$

where σ_i, $i = 1, \ldots, N_m$, denotes the standard deviation for the ith source parameter. The descent direction $\mathbf{p}$ may be pre-conditioned by $\mathbf{C}_{\boldsymbol{m}}$, such that the model update is given by

$$\delta\mathbf{m} = \lambda\, \mathbf{C_m} \cdot \mathbf{p}, \tag{22}$$

where λ denotes the step length in the descent direction.

Consider an 11-parameter source inversion involving the six elements of the moment tensor, $\{M_{11}, M_{22}, M_{33}, M_{12}, M_{13}, M_{23}\}$, the centroid location $\{x_1^s, x_2^s, x_3^s\}$, the centroid time t_s and the half duration h_s. The gradient is determined by

$$\begin{aligned}\mathbf{g} &= \{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8, g_9, g_{10}, g_{11}\} \\ &= \left\{ \frac{\partial \chi}{\partial M_{11}}, \frac{\partial \chi}{\partial M_{22}}, \frac{\partial \chi}{\partial M_{33}}, \frac{\partial \chi}{\partial M_{12}}, \frac{\partial \chi}{\partial M_{13}}, \frac{\partial \chi}{\partial M_{23}}, \frac{\partial \chi}{\partial x_1^s}, \frac{\partial \chi}{\partial x_2^s}, \frac{\partial \chi}{\partial x_3^s}, \frac{\partial \chi}{\partial t_s}, \frac{\partial \chi}{\partial h_s} \right\}. \end{aligned} \tag{23}$$

We let

$$\sigma_{1,\ldots,6} = \sigma_{\mathbf{M}} = \sqrt{2}\left(g_1^2 + g_2^2 + g_3^2 + 2g_4^2 + 2g_5^2 + 2g_6^2\right)^{-1/2}, \tag{24}$$

$$\sigma_{7,8,9} = \sigma_{\mathbf{x}_s} = \left(g_7^2 + g_8^2 + g_9^2\right)^{-1/2}, \tag{25}$$

and

$$\sigma_{10,11} = \sigma_{t_s} = \left(g_{10}^2 + g_{11}^2\right)^{-1/2}, \tag{26}$$

which guarantees that different types of CMT parameters contribute more or less equally to the magnitude of the gradient. Now we define non-dimensionalized source parameters $\hat{m}_i = m_i/\sigma_i$, such that the gradient $g_i = \partial\chi/\partial m_i$ may be replaced by the scaled gradient $\hat{g}_i = \partial\chi/\partial\hat{m}_i = g_i\,\sigma_i$.

2.2.2 Iterative scheme

At the start of the kth conjugate-gradient iteration, we have the current source model $\hat{\mathbf{m}}^k$ (a superscript denotes the iteration number) based on the previous descent direction $\hat{\mathbf{p}}^{k-1}$ and gradient $\hat{\mathbf{g}}^{k-1}$. During the kth iteration, we compute the misfit value χ and gradient $\hat{\mathbf{g}}^k$ for the current source model; we update the descent direction $\hat{\mathbf{p}}^k$ and estimate the step length in this direction, which gives us the updated model according to eq. (22). Note that estimation of step length may involve one or more additional numerical simulations, bringing the total number of simulations for each iteration to at least three. What follows is a step-by-step recipe.

(i) Non-dimensionalize the source parameters with the model covariance matrix

$$\hat{\mathbf{m}}^k = \mathbf{C}_{\mathbf{m}}^{-1/2} \cdot \mathbf{m}^k. \tag{27}$$

(ii) Calculate the value of the misfit function $\chi(\mathbf{m}^k)$, compute the gradient $\mathbf{g}^k = \partial\chi/\partial\mathbf{m}^k$ based on the adjoint method, and non-dimensionalize the gradient as $\hat{\mathbf{g}}^k = \mathbf{C}_{\mathbf{m}}^{1/2} \cdot \mathbf{g}^k$.

(iii) Determine the current descent direction $\hat{\mathbf{p}}^k = -\hat{\mathbf{g}}^k + \beta_k\,\hat{\mathbf{p}}^{k-1}$. For the first iteration, $k = 0$, set $\beta_0 = 0$ and $\hat{\mathbf{p}}^0 = -\hat{\mathbf{g}}^0$; otherwise calculate β_k based on the Fletcher–Reeves formula (Fletcher 1987)

$$\beta_k = \frac{\hat{\mathbf{g}}^k \cdot \hat{\mathbf{g}}^k}{\hat{\mathbf{g}}^{k-1} \cdot \hat{\mathbf{g}}^{k-1}} = \frac{\mathbf{g}^k \cdot \mathbf{C}_{\mathbf{m}} \cdot \mathbf{g}^k}{\mathbf{g}^{k-1} \cdot \mathbf{C}_{\mathbf{m}} \cdot \mathbf{g}^{k-1}}, \tag{28}$$

or the Pola–Ribiere formula with a direction reset (Fletcher 1987)

$$\beta_k = \max\left(0, \frac{\hat{\mathbf{g}}^k \cdot (\hat{\mathbf{g}}^k - \hat{\mathbf{g}}^{k-1})}{\hat{\mathbf{g}}^{k-1} \cdot \hat{\mathbf{g}}^{k-1}}\right), \tag{29}$$

or the Hestenes–Stiefel formula (Kelley 1999)

$$\beta_k = \frac{\hat{\mathbf{g}}^k \cdot (\hat{\mathbf{g}}^k - \hat{\mathbf{g}}^{k-1})}{\hat{\mathbf{p}}^{k-1} \cdot (\hat{\mathbf{g}}^k - \hat{\mathbf{g}}^{k-1})}. \tag{30}$$

All three formulae are equivalent for quadratic objective functions, while the Pola–Ribiere and Hestenes–Stiefel formulae may perform better for non-quadratic objective functions; we do not observe a significant improvement in convergence for our source inversions, and stick with the Fletcher–Reeves formula for our inversions.

(iv) Compute the step length λ_k in the current descent direction $\hat{\mathbf{p}}^k$. Let $f_1 = \chi(\mathbf{m}^k)$, and $g_1 = \hat{\mathbf{g}}^k \cdot \hat{\mathbf{p}}^k$.

(1) Compute a test step length $\lambda_t = -2\gamma f_1/g_1$, with a reasonable choice of γ close to 1. This assumes quadratic dependence of the misfit function on λ along the descent direction, with a minimum value of $(1-\gamma)f_1$. An equivalent formula based on the same assumption is deployed by Gauthier *et al.* (1986).

(2) Compute a test solution in the descent direction, $\hat{\mathbf{m}}_t = \hat{\mathbf{m}}^k + \lambda_t\,\hat{\mathbf{p}}^k$, and dimensionalize it: $\mathbf{m}_t = \mathbf{C}_{\mathbf{m}}^{1/2} \cdot \hat{\mathbf{m}}_t$.

(3) Compute the misfit function $\chi(\mathbf{m}_t)$ and gradient $\mathbf{g}_t = \partial\chi/\partial\mathbf{m}_t$ for subsequent cubic interpolation. Let $f_2 = \chi(\mathbf{m}_t)$, $g_2 = \hat{\mathbf{g}}_t \cdot \hat{\mathbf{p}}$, and note that we generally have $g_1 < g_2 < 0$ and $f_1 > f_2 > 0$.

(4) If quadratic interpolation is invoked, compute

$$b = [(f_2 - f_1) - g_1\lambda_t]/\lambda_t^2, \qquad c = g_1, \qquad d = f_1, \tag{31}$$

and then λ_k is given by

$$\lambda_k = \begin{cases} -c/(2b) & b \neq 0 \\ \text{error} & \text{otherwise} \end{cases} \tag{32}$$

(5) If cubic interpolation is chosen, compute

$$a = [-2(f_2 - f_1) + (g_1 + g_2)\lambda_t]/\lambda_t^3\,, \qquad b = [3(f_2 - f_1) - (2g_1 + g_2)\lambda_t]/\lambda_t^2,$$
$$c = g_1\,, \qquad d = f_1, \tag{33}$$

and then λ_k is determined by

$$\lambda_k = \begin{cases} [-b \pm (b^2 - 3ac)^{1/2}]/(3a) & a \neq 0 \text{ and } b^2 - 3ac > 0 \\ -c/(2b) & a = 0 \text{ and } b \neq 0 \\ \text{error} & \text{otherwise} \end{cases}$$

Note that when $a = 0$, cubic interpolation degenerates to quadratic interpolation. Cubic interpolation requires an extra adjoint simulation for the gradient of the test solution $\mathbf{g}_t$. We do not observe a considerable improvement in convergence with cubic interpolation, as also noted by Tape *et al.* (2007), and therefore we use quadratic interpolation in our adjoint inversions.

(v) Update the scaled source model $\hat{\mathbf{m}}^{k+1} = \hat{\mathbf{m}}^k + \lambda_k\,\hat{\mathbf{p}}^k$, and dimensionalize it: $\mathbf{m}^{k+1} = \mathbf{C}_\mathbf{m}^{1/2} \cdot \hat{\mathbf{m}}^{k+1}$.

(vi) To check convergence, examine the reduction of the misfit function for the updated model $\chi(\mathbf{m}^{k+1})$ relative to the previous model $\chi(\mathbf{m}^k)$ and the initial model $\chi(\mathbf{m}^0)$, and terminate the inversion when the target misfit reduction is achieved. In practice, we also monitor the value of $||\mathbf{g}^k||_\mathbf{m} = (\hat{\mathbf{g}}^k \cdot \hat{\mathbf{g}}^k)^{1/2} = (\mathbf{g}^k \cdot \mathbf{C}_\mathbf{m} \cdot \mathbf{g}^k)^{1/2}$, which should approach zero for the optimal model. This is equivalent to requiring $||\mathbf{m}^{k+1} - \mathbf{m}^k||_\mathbf{m} \sim 0$, that is, no further model update is necessary. Note that our objective is to iteratively reduce the misfit function, therefore even though in principle $||\mathbf{g}^k||_\mathbf{m}$ should decrease with each iteration k, there is no guarantee that $||\mathbf{g}^k||_\mathbf{m}$ decreases monotonically, depending on the topology of the non-linear misfit function in the vicinity of model $\mathbf{m}^k$.

3 TIME-REVERSAL IMAGING

The goal of this section is to develop intuition for the the adjoint wavefield $\mathbf{s}^\dagger$ and the associated adjoint strain tensor $\boldsymbol{\epsilon}^\dagger$. In the vicinity of the original source, these quantities capture the characteristics of the event. To illustrate this, consider the hypothetical, azimuthally uniform source–receiver configuration shown in Fig. 1. We use the 1-D model of Hadley & Kanamori (1977), and compute 180 s SEM seismograms generated by 10 characteristic point-source earthquakes (Fig. 2, first column) located at a depth of 7.1 km. We generate adjoint wavefields

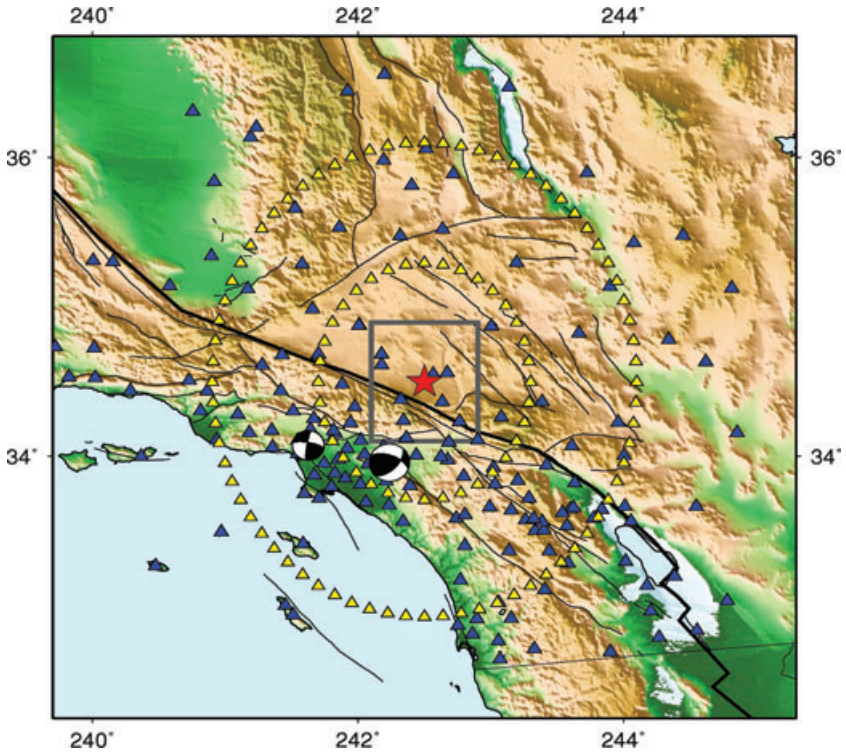

Figure 1. Hypothetical source–receiver geometry for a source located at (−117.5°, 34.5°) and labelled by the red star. Two concentric rings of receivers labelled by yellow triangles surround the hypothetical source and ensure uniform azimuthal coverage to obtain images of the source based on the adjoint wavefield $\mathbf{s}^\dagger$ and the associated adjoint strain tensor $\boldsymbol{\epsilon}^\dagger$. The grey square centred on the source denotes the close-up area shown repeatedly in Fig. 2. SCSN stations are labelled by blue triangles. Focal mechanisms for the 2001 September 9 Hollywood earthquake (depth 5.6 km) and the 2008 July 29 Chino Hills earthquake (depth 12.0 km) are plotted as beach balls. Major southern California faults are indicated by thin black lines. The North America–Pacific plate boundary is indicated by the thick black line.

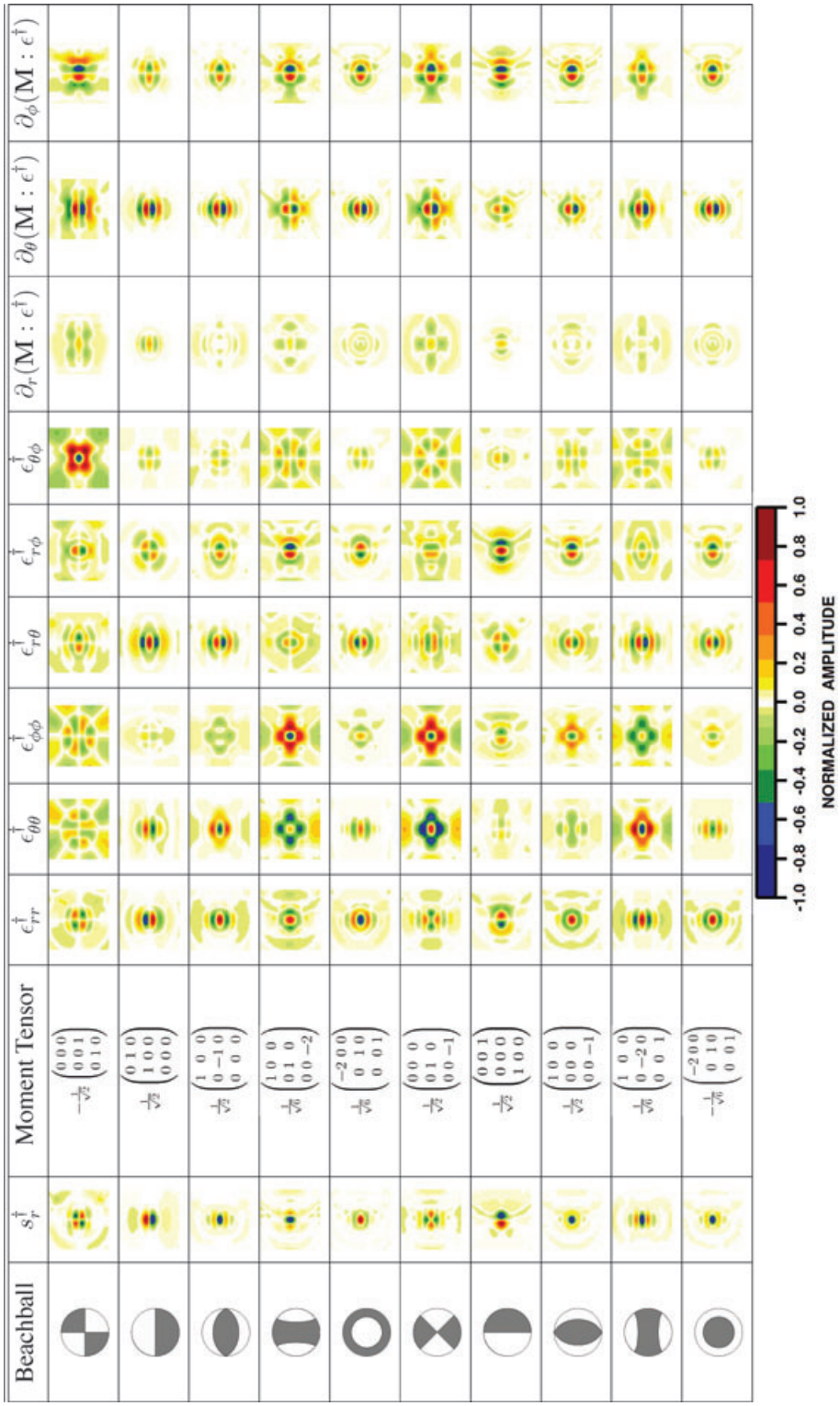

Figure 2. 10 characteristic source mechanisms specified in terms of their beach balls (first column) and moment tensors (third column) (Dahlen & Tromp 1998). Map views at the source depth and centroid time of the vertical component adjoint displacement $s_r^\dagger$ (second column), the six elements of the adjoint strain tensor $\epsilon_{rr}^\dagger$, $\epsilon_{\theta\theta}^\dagger$, $\epsilon_{\phi\phi}^\dagger$, $\epsilon_{r\theta}^\dagger$, $\epsilon_{r\phi}^\dagger$ and $\epsilon_{\theta\phi}^\dagger$, and the three centroid location Fréchet derivatives $\partial_r(\mathbf{M}:\boldsymbol{\epsilon}^\dagger)$, $\partial_\theta(\mathbf{M}:\boldsymbol{\epsilon}^\dagger)$ and $\partial_\phi(\mathbf{M}:\boldsymbol{\epsilon}^\dagger)$. The shown region is indicated by the grey square centred on the source labelled by the red star in Fig. 1. The colour scale for the vertical adjoint displacement represents the amplitude normalized by its maximum value. The colour scale for the adjoint strain-tensor elements is obtained by normalizing each amplitude by the maximum value of the six components, and that for the source mislocations by normalizing each amplitude by the maximum value of all three components.

Table 1. Synthetic inversion results for the 2001 Hollywood event. First and second row: six-parameter inversion (moment tensor). Third row: seven-parameter inversion (moment tensor plus depth). Initial solution, final solution/best inversion result based on misfit values, and true solution in order of moment magnitude, depth (km), strike (°), dip (°) and rake (°) are given in each case.

# of parameters	Initial solution	Final solution	True solution
6	4.29/7.0/270.0/60.0/20.0	4.29/7.0/280.7/73.7/6.7	4.29/7.0/281.0/74.1/7.3
6	4.29/5.0/270.0/60.0/20.0	4.18/5.0/272.9/67.6/8.9	4.29/7.0/281.0/74.1/7.3
7	4.29/5.0/270.0/60.0/20.0	4.29/6.7/280.7/74.0/6.4	4.29/7.0/281.0/74.1/7.3

$\mathbf{s}^\dagger$ by using three-component synthetic seismograms recorded by the two concentric rings of receivers as virtual sources, in accordance with eq. (9). We save the last snapshot of each adjoint wavefield to image the original source.

Fig. 2 (second column) shows map views of the vertical component of the adjoint wavefield $s_r^\dagger = \hat{\mathbf{r}} \cdot \mathbf{s}^\dagger$ at the depth and centroid time of the original source. The adjoint displacement reflects the radiation pattern near the centroid location. Fig. 2 also shows the six components of the adjoint strain tensor $\boldsymbol{\epsilon}^\dagger$ (fourth–ninth columns) and the three components of the centroid location Fréchet derivatives (10th–12th columns), as defined by eq. (12). The components of the adjoint strain tensor are correlated with the corresponding components of the original moment tensor $\mathbf{M}$, and a centroid location search should be performed in the direction indicated by the gradient of $\mathbf{M} : \boldsymbol{\epsilon}^\dagger$.

4 SYNTHETIC TEST

We test the conjugate-gradient adjoint source–inversion algorithm for two southern California earthquakes, namely, the $M_w = 4.2$ 2001 September 9, Hollywood event and the $M_w = 5.4$ 2008 July 29, Chino Hills event (Fig. 1). Both earthquakes are well recorded by broad-band SCSN seismometers and well positioned within the network. These events were previously studied based on different types of inversion schemes, including first-motion picks (Hauksson *et al.* 2002), body-wave inversions (Tan 2006), surface wave inversions (Thio & Kanamori 1995) and CMT inversions based on 3-D Green's functions (Liu *et al.* 2004). Both earthquakes were used in recent adjoint tomography of the southern California crust (Tape *et al.* 2009, 2010), showing great improvements in traveltime and waveform fits for the updated CVM-H6.2 model.

First-motion mechanisms provided by the SCSN catalogue (www.data.scec.org/catalog_search/) are used as the 'true' solutions. For most cases, we construct 'initial' moment-tensor solutions by varying the strike, dip, rake and depth of the events, and invert for the centroid-moment tensor and depth. We demonstrate nine-parameter inversions involving centroid-moment tensor and location for the Chino Hills event. Synthetics seismograms are calculated for the initial source models, and we iteratively improve estimates of source parameters based on the recipe discussed in Section 2.2.2, employing the same set of windows as is used in the actual inversion of real data described in Section 5.

4.1 Hollywood event

For the 2001 Hollywood event, we determine the six moment-tensor elements plus depth, fixing the source latitude and longitude to the 'true' epicentre. In reality, horizontal mislocations may significantly influence non-linearity of the misfit function in the vicinity of the initial solution as discussed further in Section 4.2. To begin with, we fix the event magnitude at $M_w = 4.29$ and depth at 7 km, varying only the strike, dip and rake of the event by as much as 15° (first row of Table 1). We invert for an optimal source mechanism by minimizing the misfit function within selected time windows. The first row of Fig. 3 illustrates that convergence is quickly achieved in two iterations for this six-parameter inversion, with a misfit reduction of 100 per cent. Note that each adjoint iteration involves at least two forward simulations (one for the current model and one for the test model), and one adjoint simulation (to obtain the gradient), therefore a total number of six numerical simulations is required to achieve convergence for this six-parameter inversion, which is equivalent to the number of simulations required for Hessian-based moment-tensor inversions with 3-D Green's functions (Liu *et al.* 2004). In this case, no gain is achieved by employing the adjoint inversion technique. However, when multiple moment-tensors are involved, for example, for a compound event, the number of simulations for inversions based on the adjoint method is independent of the number of sources, whereas an inversion based on 3-D Green's functions scales with the number of subevents.

Next, in addition to strike, dip and rake, we also modify the event depth from 7 km for the true solution to 5 km in our initial solution. We seek to recover the true solution based on either six- (moment tensor) or seven-parameter (moment tensor plus depth) inversions. The non-linearity of the misfit function with respect to depth traps the six parameter inversion in a local minimum, and the misfit value for the final model is even larger than for the initial model (second rows of Table 1 and Fig. 3). When depth is included, convergence is achieved in about three iterations with a misfit reduction close to 100 per cent, illustrating the importance of allowing for variations in both source mechanism and depth (third rows of Table 1 and Fig. 3). The total number of simulations adds up to $3 \times 3 = 9$ in this case, which exceeds the seven simulations involved in an inversion based on 3-D Green's functions. However, the latter approach uses only the approximate Hessian (eq. 20), and provides a solution based on a quadratic expansion of the non-linear misfit function, and more iterations may be required to achieve ultimate convergence. A more dramatic example is presented for the Chino Hills event in the next section.

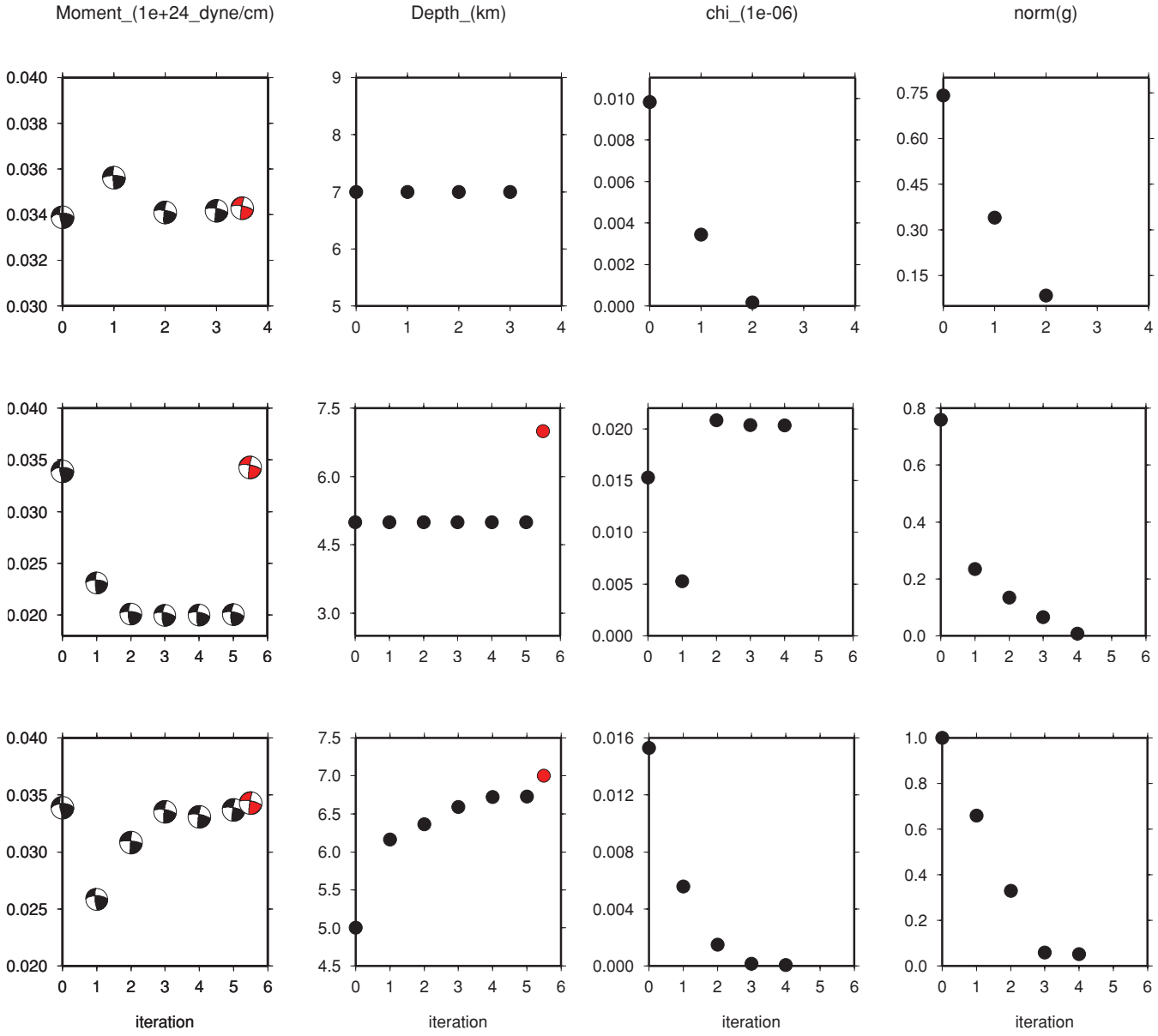

Figure 3. Synthetic inversion results for the 2001 Hollywood event. First and second row: six-parameter inversion (moment tensor). Third row: seven-parameter inversion (moment tensor plus depth). Evolution of scalar moment and source mechanism (first column), depth (second column), misfit value (third column) and magnitude of the gradient (fourth column) are shown in each case. The true mechanism and depth are denoted by a red beach ball in the first column and a red dot in the second column.

4.2 Chino Hills event

For the 2008 Chino Hills earthquake, we treat the moment-tensor solution provided by the SCSN catalogue as the 'true solution', and vary its source mechanism and depth to obtain our initial solution. The initial solution was designed to deviate significantly from the 'true' model, with differences in magnitude of 0.1, depth of up to 8 km, and strike, dip and rake of up to 14°. The seven-parameter source inversion involving source mechanism and depth based on the adjoint method converges after three iterations and gives a final solution very close to the true solution, with a misfit reduction of 98 per cent, strike, dip and rake all within $\pm$ 1°, and depth within $\pm$0.2 km (first rows of Table 2 and Fig. 4). This demonstrates the fast convergence rate for minimization of the non-linear misfit function based on conjugate-gradient techniques. The Hessian-based inversion utilizing 3-D Green's functions produces a solution with a 82 per cent misfit reduction after one iteration (third row of Table 3), thus requiring additional iterations to further reduce the misfit and resulting in an equal or larger number of simulations as the adjoint CMT method. Indeed, the second Hessian-based inversion brings the misfit reduction to almost 100 per cent (fourth row of Table 3), and gives a final solution very close to the conjugate-gradient based solution (second row of Table 3) and the true solution. Convergence requires nine simulations for the adjoint CMT method as opposed to 14 simulations for a Hessian-based inversion, making the adjoint CMT approach more computationally efficient.

Table 2. Synthetic seven-parameter inversion results for the 2008 Chino Hills event. First row: initial solution provided by the SCSN catalogue. Second row: initial mechanism and depth provided by the global CMT catalogue; centroid epicentre from the SCSN catalogue. Initial solution, final solution/best inversion result based on misfit values, and true solution in order of moment magnitude, depth (km), strike (°), dip (°) and rake (°) are given in each case.

	Initial solution	Final solution	True solution
SCSN	5.50/20.0/34.0/50.0/50.0	5.40/12.2/43.7/57.5/37.7	5.39/12.0/42.9/58.2/37.3
Global CMT	5.46/20.0/43.6/54.8/29.0	5.40/11.8/43.6/56.3/40.8	5.39/12.0/42.9/58.2/37.3

Table 3. Synthetic seven-parameter inversion results for the 2008 Chino Hills event. The first Hessian-based inversion (CMT3D) uses the same initial as in the first row of Table 2, while the second Hessian inversion uses the first CMT3D result as the initial solution. All solutions are listed in the order of moment magnitude, depth (km), strike (°), dip (°) and rake (°). Misfit reduction values of various inversion results with respect to that given by the initial SCEC source parameters are also given. The result for conjugate-gradient based inversions (CMT3D-CG) from the first row of Table 2 is listed for comparison.

	Chino Hills source parameters	Misfit reduction
SCEC (initial)	5.50/20.0/34.0/50.0/50.0	
First CMT3D inversion	5.41/14.9/44.5/61.2/32.1	82 per cent
Second CMT3D inversion	5.38/11.6/42.7/57.4/38.3	99 per cent
True solution	5.39/12.0/42.9/58.2/37.3	
CMT3D-CG	5.40/12.2/43.7/53.5/37.7	98 per cent

Next, we investigate the influence of a mislocated initial solution. We construct our initial solution by varying nine source parameters, namely, epicentre, depth and source mechanism. When the initial solution is laterally mislocated by 10 km, no convergence is achieved based on the conjugate-gradient inversions. When the lateral mislocation is reduced to 5 km, sufficient convergence is achieved in about five iterations with an error of ± 0.2 km (second row of Table 4 and third row of Fig. 4). Note that when latitude and longitude are included in the conjugate-gradient inversion, the misfit function becomes more non-linear, and more iterations are required to reach convergence.

Finally, we retrieved the CMT solution for the 2008 Chino Hills event from the global CMT catalogue (www.globalcmt.org) and use the moment tensor and depth from this catalogue in combination with the event latitude and longitude from the SCSN solution to define a new initial source model. We choose not to use the original epicentre location from global CMT solution because its epicentre differs by 9 km from the SCEC catalogue, resulting in divergence of the conjugate–gradient algorithm. The inversion results summarized in the second row of Table 2 and Fig. 4 show that full convergence is achieved in three–four iterations, with a misfit reduction of 96 per cent. As discussed in Section 2.2.2, the norm of the gradient may not decrease monochromatically due to non-quadratic variations of the misfit function in the vicinity of its global minimum.

5 ADJOINT MOMENT-TENSOR INVERSIONS

In this section we use the adjoint CMT algorithm to determine source mechanisms for the 2001 Hollywood and 2008 Chino Hills earthquakes based on three-component broad-band data recorded by the SCSN. We deconvolve the instrument response from the data in the period range from 1 to 50 s. Using initial moment-tensor solutions provided by SCSN catalogue, we compute synthetic seismograms for the SCEC CVM-H6.2 model.

For southern California earthquakes with magnitudes less than ∼5.5, half durations determined based on the relation $h_s = 2.4 \times 10^{-6}$ M_0, where M_0 denotes the scalar moment (Dahlen & Tromp 1998), are less than 0.4 s (Tape *et al.* 2010). Model CVM-H6.2 produces synthetics with reasonable fits to data at periods longer than 2 s for body waves and 5 s for surface waves. Therefore, we are unable to resolve half duration as a source model parameter. We allow for time-shifts between data and synthetics to accommodate additional unmodelled 3-D heterogeneities (Liu *et al.* 2004). Since most southern California events are well located in terms of their epicentres (Lin *et al.* 2007a,b), in what follows we generally invert for centroid-moment tensors and depth. Only for the Chino Hills event, where epicentral locations from the global CMT and SCEC catalogues differ significantly, do we consider source inversions for source mechanism, depth and epicentre.

Table 4. Synthetic nine-parameter inversion results for the 2008 Chino Hills event. Initial solution, final solution/best inversion result based on misfit values, and true solution in order of moment magnitude, latitude (°), longitude (°), depth (km), strike (°), dip (°) and rake (°) are given in each case.

	Chino Hills source parameters
Initial solution	5.50/33.986/− 117.733/20.0/43.6/54.8/29.0
Final solution	5.40/33.955/− 117.766/12.4/42.5/61.0/33.0
True solution	5.39/33.956/− 117.763/12.0/42.9/58.2/37.3

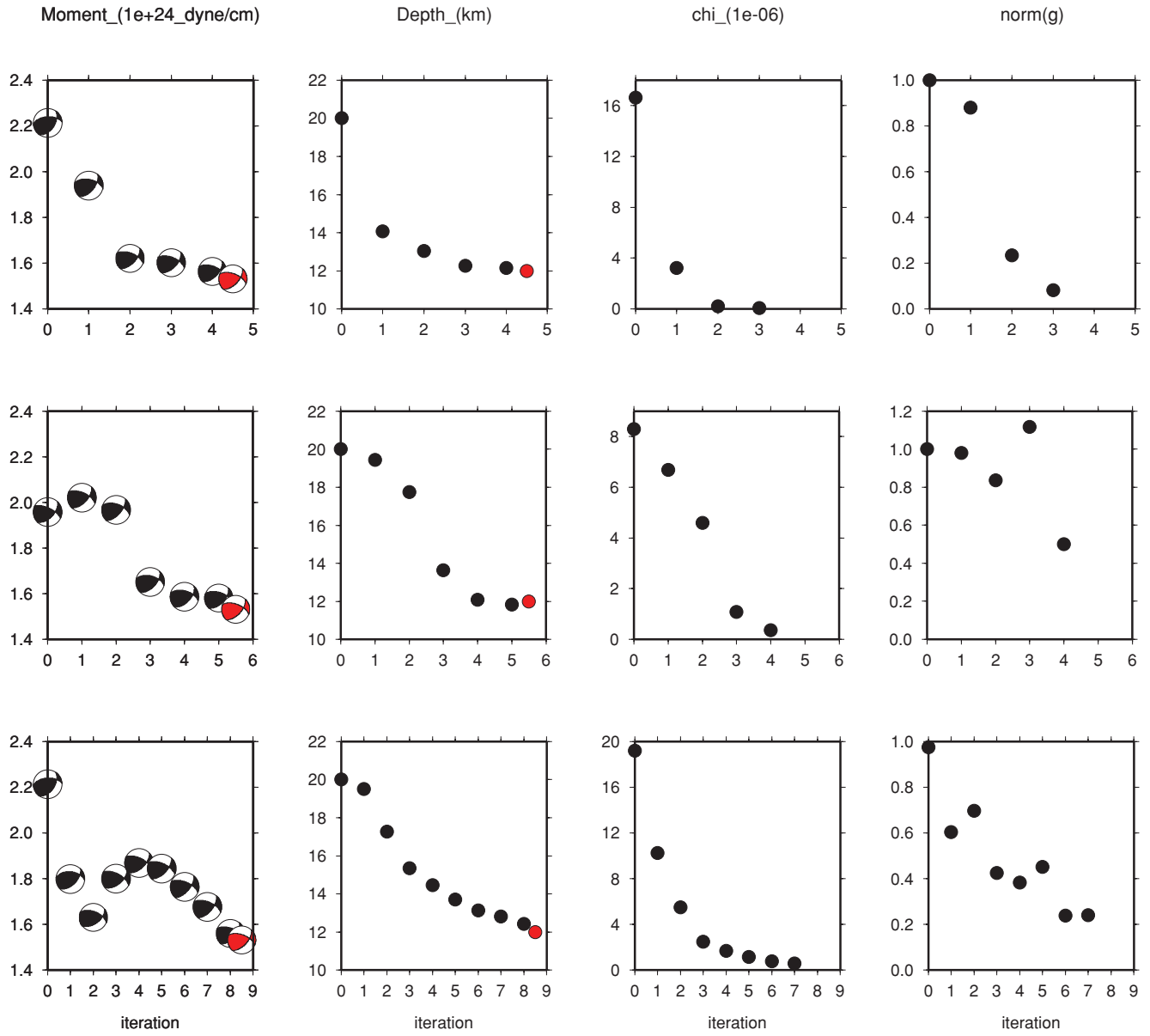

Figure 4. Synthetic inversion results for the 2008 Chino Hills event. First row: initial solution is modified from the SCSN catalogue for the seven-parameter inversion. Second row: initial mechanism and depth provided by the global CMT catalogue; centroid epicentre from the SCSN catalogue for the seven-parameter inversion. Third row: initial mechanism and hypocentre are modified from the SCSN catalogue for the nine-parameter inversion. Evolution of scalar moment and source mechanism (first column), depth (second column), misfit value (third column) and magnitude of the gradient (fourth column) are shown in each case. The true mechanism and depth are denoted by a red beach ball in the first column and a red dot in the second column.

We select time windows in which data and synthetics have relatively good waveform fits based on the FLEXWIN package (Maggi *et al.* 2009). This package calculates a reference trace based on the ratio of short-term and long-term averages of the synthetic, and from it determines seed windows containing impulsive or wavegroup-like features. Data and synthetics are compared within these seed windows based on amplitude ratios, cross-correlation values as well as cross-correlation time-shifts, and those windows with values exceeding pre-selected reference values are discarded. In this particular case, we require cross-correlation values to be larger than 0.71, amplitude ratios to be smaller than 2.5 and time-shifts to be less than 4.0 s. When records are low-pass filtered at 5 s, most selected windows contain surface waves (Rayleigh and Love waves) and long-period body waves ($P_{\rm nl}$). We employ a weighting scheme similar to that used by Liu *et al.* (2004), defining the weight of a specific window by

$$W_p = W_c \left(\frac{\Delta}{\Delta_{\rm ref}}\right)^{W_d} \frac{{\rm CC}^{W_r}}{N_p^{W_a}}, \tag{34}$$

where W_c, W_d, W_r and W_a are weights assigned to this window based on its component (W_c), distance $\Delta(W_d)$, cross-correlation value CC (W_r) and azimuth (W_a). The reference distance $\Delta_{\rm ref}$ is assigned depending on wave type, with values taken from Zhu & Helmberger (1996). The number of selected traces within the same azimuthal bin is denoted by N_p (a total of 10 bins covering a 360° azimuthal range). Therefore,

Table 5. Inversion results for the 2001 Hollywood earthquake based on SCSN data. Listed are the initial solution, final solution/best inversion result based on misfit, and Hessian-based 3-D Green's tensor solution in order of moment magnitude, depth (km), strike (°), dip (°) and rake (°).

Initial solution	Final solution	CMT3D solution
4.29/5.0/281.0/74.1/7.3	4.19/5.9/266.1/85.7/1.4	4.08/5.1/273.9/74.5/10.5

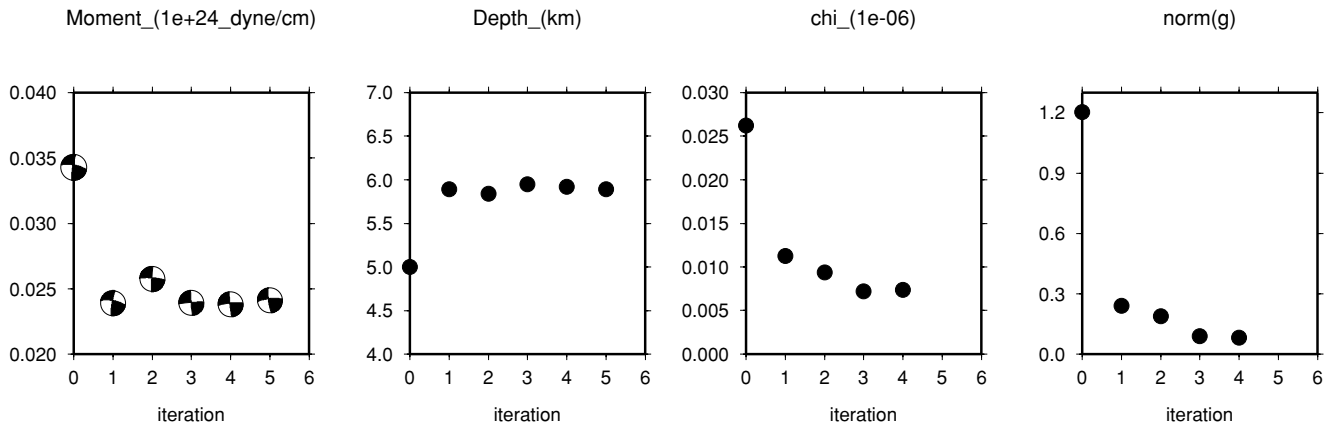

Figure 5. Inversion results for the 2001 Hollywood earthquake. Evolution of scalar moment, source mechanism, depth, misfit value and magnitude of the gradient as a function of iteration number. The inversion converges after three iterations.

the more windows are selected in a particular azimuthal bin, the less weight these windows are given in the misfit function. Individual weights are normalized with the maximum of all weights.

We apply time-shifts calculated for each window to account for unmodelled 3-D heterogeneity. We do not reselect windows unless source parameters change significantly from one iteration to the next. We evaluate the misfit function in eq. (1) for selected windows based on time-shifts and weights, and generate adjoint sources associated with the waveform difference between data and synthetics within these windows. We calculate the gradient of the misfit function with respect to the source parameters based on 3-D SEM simulations involving the adjoint eq. (9). The recipe discussed in Section 2.2 is used to determine optimal source solutions.

5.1 Hollywood event

For the Hollywood event, a total of 188 windows from 74 vertical component, 46 radial component and 68 transverse component seismograms are selected. The iterative inversion results are summarized in Table 5 and Fig. 5. Convergence is achieved within three iterations. The depth for the final solution differs from that of the initial solution by 0.9 km, which, together with the moment-tensor updates, results in a 72 per cent misfit reduction. By comparison, a Hessian-based source inversion utilizing the same initial solution and the same set of windows produces a solution with a similar misfit reduction of 68 per cent. A second Hessian-based inversion does not reduce the misfit function much further, reflecting the flat nature of the misfit function around the true solution due to systematic errors associated with unmodelled structural heterogeneities. Unlike the Hessian-based inversion, it is generally difficult to assign error bars to source parameters inverted based on adjoint methods, involving only first-order derivatives of the misfit function. One option is to examine variations in the source parameters during the last few iterations before convergence, assigning these variations as errors. Here we simply use the difference between the true solution and final solution in the synthetic test (Section 4) to assign error bars. For the Hollywood event, this implies a ± 0.3 km error in depth and less than $\pm 1^\circ$ errors in strike, dip and rake. Note that the estimated error bars may be significantly smaller than the actual differences between the true and inverted solutions, as the misfit function may contain a large component due to unmodelled 3-D heterogeneity (i.e. systematic errors). The assigned error bars only give a qualitative estimate of the accuracy of the inversion scheme.

5.2 Chino Hills event

For the Chino Hills event, due to its larger magnitude, a total of 344 windows is selected from 123 vertical component, 113 radial component and 108 transverse component seismograms. The adjoint CMT inversion results based on the SCSN initial solution are shown in the first rows of Table 6 and Fig. 6. This event is well recorded with excellent azimuthal coverage by broad-band stations in southern California, and synthetics for the SCSN initial solution already fit the recorded data in the selected windows quite well, such that the inversion only achieves a misfit reduction of 25 per cent, as convergence is reached in three or four iterations. Error bars assigned to the source parameters based on the synthetic tests in Section 4 are ± 0.2 km in depth, $\pm 1^\circ$ in strike, $\pm 2^\circ$ in dip and $\pm 3^\circ$ in rake.

Table 6. Inversion results for the 2008 Chino Hills earthquake based on SCSN data. First row: SCSN initial solution. Second row: initial mechanism and depth provided by the global CMT catalogue; centroid epicentre from the SCSN catalogue. Listed are the initial solution, final solution/best inversion result based on misfit in order of moment magnitude, depth (km), strike (°), dip (°) and rake (°), as well as misfit reduction values.

	Initial solution	Final solution	Misfit reduction
SCSN	5.39/12.0/42.9/58.2/37.3	5.42/15.1/44.1/56.5/36.8	20 per cent
Global CMT	5.46/20.0/43.6/54.8/29.0	5.44/17.0/47.9/50.9/37.2	18 per cent

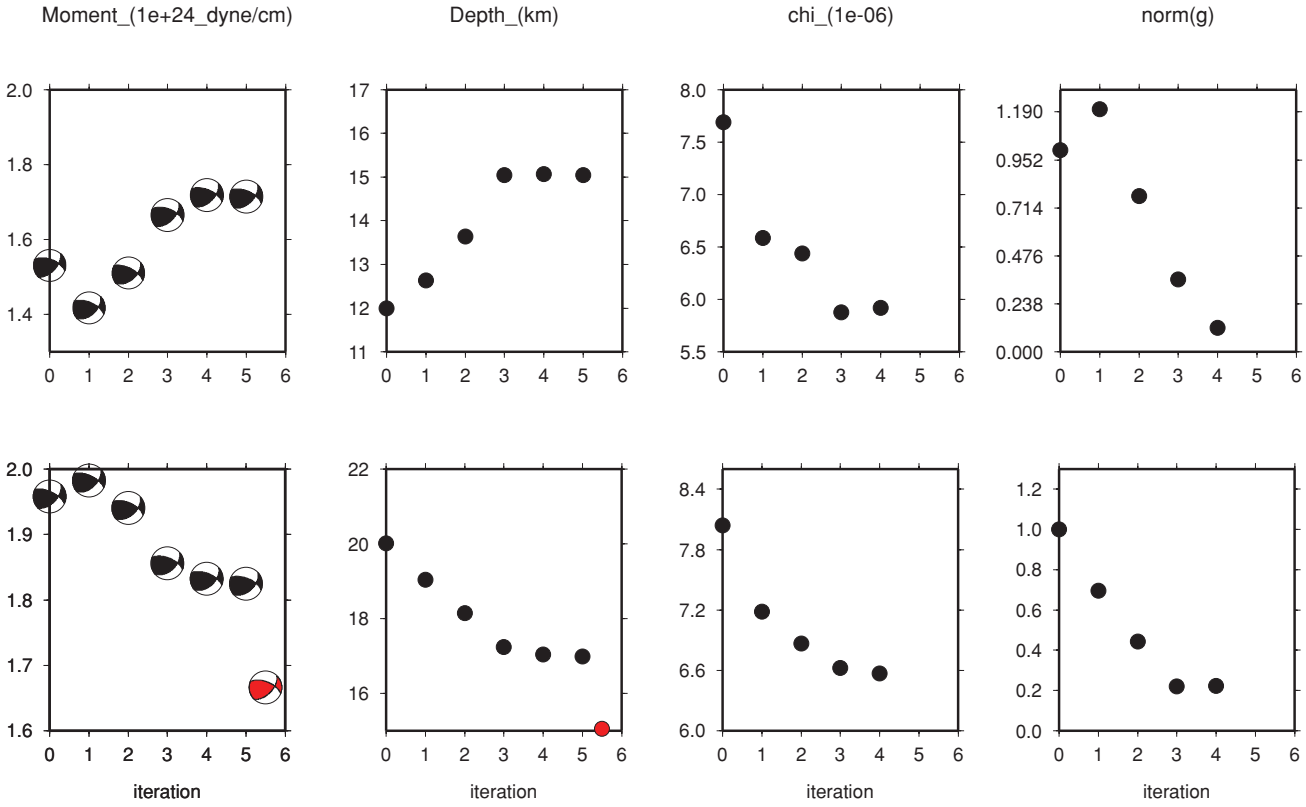

Figure 6. Inversion results for the 2008 Chino Hills earthquake. First row: SCSN initial solution. Second row: initial mechanism and depth provided by the global CMT catalogue; centroid epicentre from the SCSN catalogue. Evolution of scalar moment and source mechanism (first column), depth (second column), misfit value (third column) and magnitude of the gradient (fourth column) as a function of iteration number. The inversion converges after three or four iterations. The red beach ball in the second row represents the final solution from the first row for comparison.

We perform a second adjoint CMT inversion in which we use an initial solution from the global CMT catalogue. The event depth, source mechanism and moment magnitude of the global CMT solution are very similar to those of the SCSN solution, however, the centroid epicentre of the global CMT solution differs from the SCSN solution by about 0.08° (9 km), typical of location errors for global CMT inversions (Hjörleifsdóttir & Ekström 2010). No convergence is observed after five conjugate-gradient iterations, and the misfit function oscillates from iteration to iteration. Even when a nine-parameter inversion involving also epicentral location is performed, no convergence can be achieved for this event. This illustrates the non-linearity of the source inversion problem involving hypocentre location, and underscores the importance of a relative good initial centroid epicentre location for the recovery of other source parameters. With improvement of 3-D background velocity models, this situation may potentially be remedied for regional and global earthquakes by joint inversions of event mechanism and depth together with epicentral location, as time-shifts are no longer necessary to align data and synthetics in measurement windows, and event epicentral locations become more resolvable.

Next, we combine the global CMT source mechanism and depth with the SCSN source latitude and longitude to produce a third initial solution. With this initial solution the adjoint inversion converges in three iterations, but results in a depth of 17 km, 2 km deeper than the previously obtained solution (Fig. 6, second row). The misfit reduction is ∼18 per cent, similar to the level of misfit reduction based on the SCSN initial solution. In synthetic tests we have no difficulty in obtaining a convergent solution starting from either initial solution, but when systematic errors are large relative to the misfit associated with the source parameters, the misfit function may become highly non-linear, involving multiple local minima in which the linearized inversion may get trapped. Another possible reason for the different final solutions is that the surface wave windows selected for this inversion are relatively insensitive to depths greater than 20 km, resulting in a very flat misfit function at the optimal point.

5.3 Discussion

Most southern California earthquakes with $M_w \geq 3.5$ are well recorded by the SCSN, and, given an accurate 3-D structural model, a significant number of measurement windows is generally selected. Under these circumstances, source inversion problems are overdetermined. However, in regions with sparse network coverage, source parameters may be poorly constrained. In this case, a Hessian-based inversion may require damping to stably invert a Hessian matrix with a large condition number, while a conjugate-gradient-based adjoint source inversion does not suffer from instability. Iterative flattening of the misfit function indicates that, given the data coverage, the 'best' source model has been determined. However, this 'best' source model may depend on the chosen initial model.

6 CONCLUSIONS

We determine centroid-moment tensor solutions based on an adjoint method in combination with a non-linear conjugate–gradient algorithm. We demonstrate the applicability of the method for the 2001 Hollywood and 2008 Chino Hills earthquakes, using the recent SCEC CVM-H6.2 southern California crustal model. In general, the computational cost of the adjoint CMT algorithm is less than or comparable to the traditional Hessian-based CMT inversion utilizing 3-D Green's functions. This study prepares the way for multiple CMT and kinematic source inversions based on 3-D background models, where the adjoint CMT method is preferable, in particular in the kinematic case, because it does not presume a given fault location or orientation, making it currently the only computationally viable method. One disadvantage of the adjoint CMT method compared to an inversion based on 3-D Green's functions is that the adjoint approach is iterative, and therefore forward and adjoint simulations have to be launched sequentially, and the next iteration cannot start until the current iteration is complete. In contrast, the Hessian-based inversion permits the computation of all 3-D Green's functions simultaneously, making it suitable for parallel job submission. However, as we illustrated for the 2008 Chino Hills event, inversions for centroid location are non-linear and require an iterative approach, and such iterations must also be carried out sequentially.

ACKNOWLEDGMENTS

The numerical simulations for the time-reversal images of the source were performed on the Caltech Division of Geological & Planetary Sciences Dell cluster. SEM simulations for the moment-tensor inversions are accomplished on the tightly coupled system (TCS) component of the Scinet Consortium. SciNet is funded by the Canada Foundation for Innovation (CFI) and by the Province of Ontario as well as the University of Toronto, Faculties of Arts and Science, Engineering and Medicine. Data and synthetic seismogram processing is accomplished based on the Seismic Analysis Code (SAC, Goldstein & Snoke 2003). All plots are made with the Generic Mapping Tool (GMT) (Wessel & Smith 1991). The open source spectral-element software package SPECFEM3D used for this paper is freely available for download via the Computational Infrastructure for Geodynamics (geodynamics.org). QL is supported by the Discovery Grants of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Toronto Startup Fund. JT acknowledges support by the National Science Foundation under grant EAR-0711177. The authors thank Carl Tape for providing updated 3-D southern California model CVM-H6.2. Finally, we thank Editor Wolfgang Friederich, Yann Capdeville and an anonymous reviewer for helpful comments which improved the manuscript.

REFERENCES

Chen, P., Zhao, L. & Jordan, T.H., 2005. Finite-moment tensor of the 3 September 2002 Yorba Linda earthquake, *Bull. seism. Soc. Am.,* **95**(3), 1170–1180.

Chen, P., Jordan, T.H. & Zhao, L., 2010. Resolving fault plane ambiguity for small earthquakes, *Geophys. J. Int.,* **181**(1), 493–501.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton University Press, Princeton, NJ.

Dreger, D. & Helmberger, D.V., 1991. Source parameters of the Sierra Madre earthquake from regional and local body waves, *Geophys. Res. Lett.,* **18**(11), 2015–2018.

Dziewonski, A.M., Chou, T.A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Ekström, G., 2007. Global seismicity : results from systematic waveform analyses, 1976-2005, in *Treatise of Geophysics,* pp. 473–481, ed. Kanamori H., Elsevier, Amsterdam.

Fletcher, R., 1987. *Practical Methods of Optimization,* 2nd edn, John Wiley & Sons, New York, NY.

Gauthier, O., Virieux, J. & Tarantola, A., 1986. Two-dimensional nonlinear inversion of seismic waveforms: numerical results, *Geophysics,* **51**(7), 1387–1403.

Goldstein, P. & Snoke, A., 2003. SAC availability for the IRIS community, *DMS Electronic Newsletter,* **VII**(1). Available online at: http://www.iris.edu/news/newsletter/vol7no1/page1.htm

Hadley, D. & Kanamori, H., 1977. Seismic structure of the Transverse Ranges, California, *Geol. Soc. Am. Bull.,* **88**(71009), 1469–1478.

Hartzell, S.H. & Heaton, T.H., 1983. Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, Earthquake, *Bull. seism. Soc. Am.,* **73**(6), 1553–1583.

Hartzell, S., Liu, P.C. & Mendoza, C., 1996. The 1994 Northridge, California, earthquake: investigation of rupture velocity, risetime, and high-frequency radiation, *J. geophys. Res.,* **101**(B9), 20 091–20 108.

Hauksson, E., 2000. Crustal structure and seismicity distribution adjacent to the Pacific and North America plate boundary in southern California, *J. geophys. Res.,* **105**(B6), 13 875–13 903.

Hauksson, E., Jones, S.P. & Hutton, K., 2002. Emerging from the stress shadow of the 1992 Mw 7.3 Landers southern California earthquake? A preliminary assessment, *Seism. Res. Lett.,* **73,** 33–38.

Hauksson, E. *et al.*, 2008. Preliminary report on the 29 July 2008 Mw 5.4 Chino Hills, Eastern Los Angeles basin, California, earthquake sequence, *Seism. Res. Lett.,* **79**(6), 855–866.

Helmberger, D.V. & Engen, G.R., 1980. Modeling the long-period body waves from shallow earthquakes at regional ranges, *Bull. seism. Soc. Am.,* **70**(5), 1699–1714.

Hjörleifsdóttir, V., 2007. Earthquake source characterization using 3D numerical modeling, *PhD thesis,* Caltech, Pasadena.

Hjörleifsdóttir, V. & Ekström, G., 2010. Effects of three-dimensional Earth structure on CMT earthquake parameters, *Phys. Earth planet. Int.,* **179**(3-4), 178–190.

Hjörleifsdóttir, V., Kanamori, H. & Tromp, J., 2009. Modeling 3-D wave propagation and finite slip for the 1998 Balleny Islands earthquake, *J. geophys. Res.,* **114**(B03301), 1–15.

Ishii, M., Shearer, P.M., Houston, H. & Vidale, J.E., 2005. Extent, duration and speed of the 2004 Sumatra-Andaman earthquake imaged by the Hi-Net array., *Nature,* **435,** doi:10.1038/nature03675.

Ji, C., Wald, D.J. & Helmberger, D.V., 2002. Source description of the 1999 Hector Mine, California, earthquake, part II: complexity of slip history, *Bull. seism. Soc. Am.,* **92**(4), 1208–1226.

Kawakatsu, H. & Montagner, J.P., 2008. Time-reversal seismic-source imaging and moment-tensor inversion, *Geophys. J. Int.,* **175**(2), 686–688.

Kelley, C.T., 1999. *Iterative Methods for Optimization,* SIAM, Philadelphia, PA.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.,* **139**(3), 806–822.

Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.,* **149**(2), 390–412.

Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.,* **150,** 303–318.

Komatitsch, D., Liu, Q., Tromp, J., Süss, M.P., Stidham, C. & Shaw, J.H., 2004. Simulations of ground motion in the Los Angeles Basin based upon the Spectral-element method, *Bull. seism. Soc. Am.,* **94**(1), 187–206.

Larmat, C., Montagner, J.P., Fink, M., Capdeville, Y., Tourin, A. & Clévédé, E., 2006. Time-reversal imaging of seismic sources and application to the great Sumatra earthquake, *Geophys. Res. Lett.,* **33**(L19312), doi:10.1029/2006GL026336.

Larmat, C., Liu, Q., Tromp, J. & Montagner, J.P., 2008. Time reversal location of glacial earthquakes, *J. geophys. Res.,* **113**(B09314), doi:10.1029/2008JB005607.

Larmat, C., Guyer, R.A. & Johnson, P.A., 2010. Time-reversal methods in geophysics, *Physics Today,* **63,** 31–35.

Lin, G.Q., Shearer, P.M. & Hauksson, E., 2007a. Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *J. geophys. Res.,* **112**(B12), doi:10.1029/2007JB004986.

Lin, G.Q., Shearer, P.M., Hauksson, E. & Thurber, C.H., 2007b. A three-dimensional crustal seismic velocity model for southern California from a composite event method, *J. geophys. Res.,* **112**(B11), doi:10.1029/2007JB04977.

Lin, G.Q., Thurber, C.H., Zhang, H., Hauksson, E., Shearer, P.M., Waldhauser, F., Brocher, T.M. & Hardebeck, J., 2010. A California statewide three-dimensional seismic velocity model from both absolute and differential times, *Bull. seism. Soc. Am.,* **100**(1), 225–240.

Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. seism. Soc. Am.,* **96**(6), 2283–2397, doi:10.1785/0120060041.

Liu, Q., Tromp, J., Polet, J. & Komatitsch, D., 2004. Spectral-element moment tensor inversions for earthquakes in southern California, *Bull. seism. Soc. Am.,* **94**(5), 1748–1761.

Lovely, P., Shaw, J.H., Liu, Q. & Tromp, J., 2006. A structural Vp model of the Salton Trough, California, and its implications for seismic hazard, *Bull. seism. Soc. Am.,* **96**(5), 1882–1896.

Maggi, A., Tape, C.H., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.,* **178**(1), 257–281.

Olson, A.H. & Aspel, R.J., 1982. Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake, *Bull. seism. Soc. Am.,* **72**(6), 1969–2001.

Olsen, K.B. *et al.*, 2006. Strong shaking in Los Angeles expected from southern San Andreas earthquake, *Geophys. Res. Lett.,* **33**(7), doi:10.1029/2005GL025472.

Plesch, A., Tape, C.H. & Shaw, J.H., 2009. CVM-6.0: Inversion integration, the San Joaquin valley and other advances in the community velocity model, in *SCEC Annual Meeting, Proceedings and Abstracts,* pp. 260–261.

Ritsema, J. & Lay, T., 1993. Rapid source mechanism determination of large earthquakes in the western United States, *Geophys. Res. Lett,* **20**(15), 1611–1614.

Süss, M.P. & Shaw, J.H., 2003. P wave seismic velocity structure derived from sonic logs and industry reflection data in the Los Angeles basin, California, *J. geophys. Res.,* **108**(B3, 2170), doi:10.1029/2001JB001628.

Tan, Y., 2006. Broadband waveform modeling over a dense seismic network, *PhD thesis,* Caltech, Pasadena.

Tape, C.H., Liu, Q. & Tromp, J., 2007. Finite-frequency tomography using adjoint methods - Methodology and examples using membrane surface waves, *Geophys. J. Int.,* **168,** 1105–1129.

Tape, C.H., Liu, Q., Tromp, J. & Maggi, A., 2009. Adjoint tomography of the southern California crust, *Science,* **325**(5943), 988–992.

Tape, C.H., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180**(1), 433–462.

Thio, H.K. & Kanamori, H., 1995. Moment-tensor inversions for local earthquakes using surface waves recorded at TERRAscope, *Bull. seism. Soc. Am.,* **85**(4), 1021–1038.

Tromp, J., Tape, C.H. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160,** 195–216.

Walker, K.T., Ishii, M. & Shearer, P.M., 2005. Rupture details of the 28 March 2005 Sumatra Mw 8.6 earthquake imaged with teleseismic P waves, *Geophys. Res. Lett.,* **32**(24), doi:10.1029/2005GL024395.

Wallace, T.C., Helmberger, D.V. & Mellman, G.R., 1981. A technique for the inversion of regional data in source parameter studies, *J. geophys. Res.,* **86**(B3), 1679–1685.

Wessel, P. & Smith, W.H.F., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.,* **72**(41), 441.

Zhao, L.S. & Helmberger, D.V., 1994. Source estimation from broadband regional seismograms, *Bull. seism. Soc. Am.,* **84**(1), 91–104.

Zhu, L. & Helmberger, D.V., 1996. Advancement in source estimation techniques using broadband regional seismograms, *Bull. seism. Soc. Am.,* **86**(5), 1634–1641.

Zhu, L. & Kanamori, H., 2000. Moho depth variation in southern California from teleseismic receiver functions, *J. geophys. Res.,* **105**(B2), 2969–2980.